



Research article

Enhancing traditional Chinese medicine diagnostics: Integrating ontological knowledge for multi-label symptom entity classification

Hangle Hu¹, Chunlei Cheng^{1,2,*}, Qing Ye¹, Lin Peng¹ and Youzhi Shen¹

¹ School of Computer Science, Jiangxi University of Chinese Medicine, Nanchang 330004, China

² Key Laboratory of Artificial Intelligence in Chinese Medicine, Jiangxi University of Chinese Medicine, Nanchang 330004, China

* **Correspondence:** Email: 20060870@jxutcm.edu.cn, huhangle@jxutcm.edu.cn; Tel: +8613870990756.

Abstract: In traditional Chinese medicine (TCM), artificial intelligence (AI)-assisted syndrome differentiation and disease diagnoses primarily confront the challenges of accurate symptom identification and classification. This study introduces a multi-label entity extraction model grounded in TCM symptom ontology, specifically designed to address the limitations of existing entity recognition models characterized by limited label spaces and an insufficient integration of domain knowledge. This model synergizes a knowledge graph with the TCM symptom ontology framework to facilitate a standardized symptom classification system and enrich it with domain-specific knowledge. It innovatively merges the conventional bidirectional encoder representations from transformers (BERT) + bidirectional long short-term memory (Bi-LSTM) + conditional random fields (CRF) entity recognition methodology with a multi-label classification strategy, thereby adeptly navigating the intricate label interdependencies in the textual data. Introducing a multi-associative feature fusion module is a significant advancement, thereby enabling the extraction of pivotal entity features while discerning the interrelations among diverse categorical labels. The experimental outcomes affirm the model's superior performance in multi-label symptom extraction and substantially elevates the efficiency and accuracy. This advancement robustly underpins research in TCM syndrome differentiation and disease diagnoses.

Keywords: biomedical system analysis; medical informatics; multi-label entity extraction; natural language processing; artificial intelligence

1. Introduction

In natural language processing (NLP), named entity recognition (NER) is a cornerstone technology, particularly vital in traditional Chinese medicine (TCM). It focuses on the automated recognition and classification of entities with specific meanings within the text. In TCM, this technology is instrumental in annotating and classifying symptoms, facilitating symptom standardization [1], and plays a crucial role in classifying symptoms [2], categorizing relationships [3,4], and constructing knowledge graphs [5]. These advancements subsequently bolster support for auxiliary diagnostics [6,7].

However, the application of NER in TCM faces limitations due to the requisite for standardized categorization in TCM clinical entities [8–11] and named entities within TCM literature [12]. These texts often present challenges such as complex nomenclatures, ambiguous definitions, and arbitrary combinations of compound symptoms [13]. Despite recent advancements in the standardization of TCM symptoms [14] and information extraction about TCM symptoms [15]—emphasizing terminological and conceptual normalization [16,17], symptom classification and grading [18], standardized data collection, and clinical diagnostic relevance [19]—the existing public ontology sets remain inadequate. Recent research initiatives, such as Ma et al. [9], developed multi-granularity text feature encoders tailored for TCM literature’s NER. Similarly, Qi et al. [10] introduced an enhanced Tri-Training semi-supervised learning algorithm and a multi-neural network fusion model which significantly boosted the TCM entity recognition task performance, especially under the constraints of limited labeled data. However, these methods for extracting TCM entities are primarily restricted to a limited range of categories such as diseases, diagnoses, and medications, and most need to integrate domain-specific knowledge, thus curbing their overall effectiveness. The need for a comprehensive and scientifically robust TCM ontology framework is paramount. Such a framework would be a benchmark for a fine-grained classification and integrate structured TCM knowledge into models, thereby enhancing the accuracy and efficiency.

A fine-grained entity type classification is crucial for accurately identifying semantic categories of entities in unstructured text [20,21], particularly in extracting symptom entities, constructing narratives, and retrieving critical information to reflect the complexity and specificity of conditions in TCM records. Ren et al. [22] proposed a two-stage approach for cleaning and leveraging distantly labeled data. Onoe et al. [23] introduced a probabilistic auto-relabelling method to address noise in fine-grained, entity-type recognition under distant supervision. Furthermore, Zhang et al. [24] enhanced the representation in noisy named entity type recognition through edge-weighted attention graph convolutional networks. Lastly, Ali et al. [25] presented a novel embedding framework to reduce label noise in entity-type recognition based on distant supervision. These methods transform fine-grained entity-type classification into a multi-label classification problem, aim to identify specific entities in text, predict their likelihood scores, and classify them based on either these scores or hierarchical-type structures. However, these approaches grapple with challenges such as category imbalance and the complexity of type hierarchies, often resulting in algorithms overfitting common categories, escalating inference difficulties, and increasing computational demands, sometimes leading to classification ambiguities. Advanced multi-label classification methodologies have been introduced to counter these issues. Prabhu et al. [26] enhanced scalability and maintained accuracy by developing a balanced label hierarchy. You et al. [27] presented an efficacious approach which utilized a multi-label attention mechanism coupled with a probabilistic label tree, thereby significantly improving the handling of

extensive label datasets. Zhang et al. [28] implemented a recursive fine-tuning strategy, thereby substantially expediting the training process within large label dimensions. These methods enhance the efficiency of large-scale label classification by semantically grouping many labels into meta-labels and utilizing hierarchical label, tree-based, multi-stage shortlisting for effective recursive clustering.

Nonetheless, these multi-label classification strategies primarily focus on text classification, with the principal objective of assigning multiple labels to each input instance rather than directly extracting the entity information from texts. In applying these strategies to entity recognition tasks, there is a necessity to supplement them with entity-specific features. Our methodology integrates entity recognition models with multi-label classification techniques to extract essential entity information from texts. This integration effectively manages the diversity of entity categories and the intricacy of label structures. For instance, in analyzing a TCM diagnostic case such as “Present Illness History,” symptom entities such as “difficulty in falling asleep,” “frequent dreaming,” and “easily awakened” are extracted from relevant texts. The identified entities are converted into feature representations and, combined with contextual embeddings, are fed into subsequent multi-label classification modules. This process demonstrates the effective integration of entity recognition and multi-label classification in analyzing TCM texts, thus highlighting their complementary roles in enhancing the text analysis accuracy.

This study offers a sophisticated approach to symptom entity recognition and classification within the field of TCM diagnostics, thereby proposing a multi-layered solution strategy. Initially, we establish a comprehensive TCM symptom ontology framework based on authoritative TCM literature, thereby integrating domain knowledge through knowledge graphs to enhance the accuracy of symptom descriptions and classifications. Subsequently, we effectively address the diversity of entity categories and the vastness of the label space by employing a two-stage entity classification process in tandem with multi-label classification methods. Finally, we have designed a correlated feature fusion module (CFFM) that amalgamates attention mechanisms and multi-layer perceptrons. This module optimizes the entity information extracted by the recognition model for multi-label classification and captures interdependencies between different category labels, thereby enhancing the model’s feature representation capabilities. Our contributions can be summarized as follows:

- 1) The construction of an all-encompassing and precise TCM symptom ontology framework, which integrates multi-level classifications and four key dimensions, and is anchored in authoritative texts to ensure its high reliability and scientific validity.

- 2) The implementation of our method strategically combines TCM knowledge with the model by merging the ontology framework with the model through knowledge graphs. This integration ensures a seamless fusion of domain-specific insights and model functionalities. This approach ensures that the model is driven by knowledge and dynamically updates the ontology framework in response to changes in text data.

- 3) The development of the CFFM optimized for refining entity features extracted by the recognition model, thus enhancing the multi-label classification tasks. Moreover, by integrating knowledge graph embedding techniques, the CFFM efficiently refines inter-label correlation features, promoting the quality of the hierarchical label tree structure and significantly bolstering the model’s capacity to process complex label systems.

2. Related works

In the last two decades, substantial progress has been made in the standardization of symptom terminology and classification within TCM [14]. Xiao et al. [29] applied machine learning techniques to extract symptom terms from over 1,900 clinical records, using technologies such as hidden markov models (HMM) and conditional random fields (CRF) to ensure comprehensive symptom descriptions. Zhang et al. [30] compiled diagnostic and symptomatic terms for epilepsy, thus establishing a symptom corpus and employing methods such as core symptom extraction to normalize terminology. Drawing from TCM texts, Li et al. [31] refined and organized symptom terms by developing a symptom lexicon. This study adopts analogous methodologies for TCM symptoms to guarantee the reliability of terminological standards, thereby systematically categorizing and classifying symptoms to affirm their thoroughness and integrity.

In the sphere of NLP, traditional research areas like NER [32] and entity typing [33] have recently evolved to focus on fine-grained entity typing (FET) and ultra-fine entity typing (UFET) [34], aimed at accurately predicting the specific subtypes of entities. A significant challenge in this field is the effective management of hierarchical ontologies. Prior research has predominantly treated hierarchical typing as a multi-label classification task, thereby incorporating hierarchical structures in various ways and endeavoring to unearth more extensive label information [35,36] or to enhance label representation [37]. In multi-label classification, the typical selection of feature subsets encompassing label set information involves using filters, wrappers, and embeddings [38–41]. For managing extensive label quantities, these methods usually utilize filter methods for feature extraction, most notably constructing hierarchical label trees (HLT) using TF-IDF [28,42]. This strategy facilitates feature selection prior to model training, thereby significantly reducing the computational overhead and enhancing the model's interpretability.

However, these methods encounter distinct challenges when applied to the specialized and intricate domain of TCM diagnostic texts, such as recognizing specific terminologies, navigating linguistic complexities, and overcoming data sparsity. Our research addresses these issues by implementing a CFFM integrated with domain-specific knowledge embeddings to effectively capture and interpret label-related features. Furthermore, our study extends the scope of multi-label classification in TCM texts, by focusing on assigning multiple labels to each input example. We utilize entity recognition methods to capture the vital entity information and integrate this with TCM knowledge for a more refined and detailed entity classification, thereby contributing to the precision and depth of TCM symptom identification.

3. Materials and methods

This research presents a sophisticated model comprised of three pivotal modules, as shown in Figure 1: the entity recognition module, the correlation feature fusion module, and the multi-label classification module. The exposition of this study commences with an articulate presentation of the TCM symptom ontology framework's architecture, subsequently delving into an in-depth exploration of the functionalities and intricacies of each module.

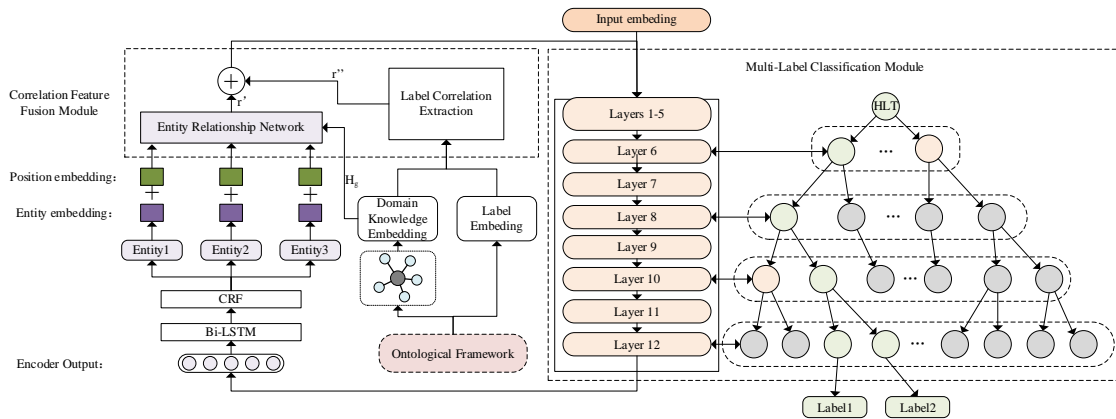


Figure 1. Overall architecture.

3.1. Traditional Chinese medicine symptom ontology framework

This study aims to standardize TCM terminology by integrating diverse symptom descriptions derived from authoritative TCM texts, thus forming a comprehensive knowledge system. It categorizes symptoms from four distinct perspectives, as illustrated in Figure 2: “origin (new compilation of TCM diagnoses),” “disease (TCM et al.),” “observation (standard terms of common clinical symptoms in TCM),” and “specialty (practical training on four diagnostic skills of TCM).”

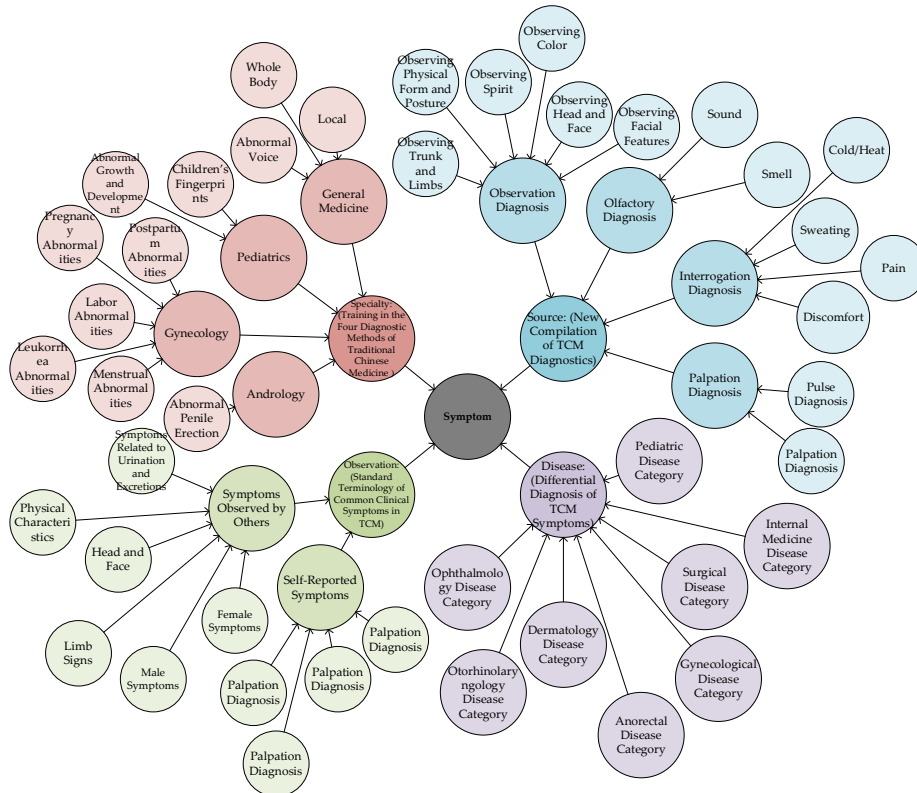


Figure 2. Visual knowledge graph of traditional Chinese medicine symptom ontology framework.

We adopted a methodology from Sun et al. [43] to construct the symptom ontology framework and executed it in three stages. Initially, we excluded catalogs describing diagnostic methods and established a top-down categorization. Subsequently, we applied a bottom-up clustering approach rooted in the symptom terminology and the underlying knowledge system. Ultimately, we consolidated the findings to finalize the ontology categories. Given our team's limited expertise in TCM, we diligently tried to preserve the original symptom names and classifications from the monographs, thereby ensuring the framework's scientific rigor and validity. We meticulously edited the ontology using the Protege tool, which supports the Chinese language and is user-friendly, thereby offering detailed statistical insights into the various entity categories within the ontology structure (see Table 1).

Table 1. The specific symptom information within the ontology architecture.

	Source	Disease	Observation	Specialty	Total
Number of categories	27	19	37	36	119
Number of symptom entities	1056	684	840	923	3503

3.2. Entity recognition module

Currently, mainstream entity recognition models can only classify each entity separately. However, achieving multi-label entity classification in text requires more entity-related information. Therefore, we need to extract accurate and effective entity features to improve the efficiency and accuracy of multi-label entity extraction. For this purpose, to input sequences of TCM medical records $X = \{x_1, x_2, \dots, x_n\}$, we extract entities E , entity embeddings E_i , and entity position encoding p_i from the BERT + Bi-LSTM + CRF model. We obtain entity boundaries (B_e, E_e) through sequence labeling, utilize BERT embeddings $(BERT(E_i))$ to capture the semantic information of entities v_i , and compute entity position encoding $p_i = f(B_e, E_e)$ to specify their scope within the text.

We utilize CRF to model the entity extraction, and the loss function incorporates the log-likelihood loss of CRF. The loss function can be represented as follows:

$$L_{entity} = -\frac{1}{N} \sum_{i=1}^N \log(P(y_i|x_i; \theta)) + \lambda \cdot \|\theta\|^2 \quad (1)$$

In this context, θ represents the model parameters, $P(y_i|x_i; \theta)$ stands for the predicted label sequence, λ signifies the regularization parameter, and $\|\theta\|^2$ corresponds to the square of the L2 norm of the model parameters.

3.3. Correlation feature fusion module

TCM perceives diseases as holistic entities, where symptoms and signs are intricately interconnected and reflect multifaceted aspects of the disease; thus, discerning the interrelationships between entities is paramount. To generate feature representations of entity information and to focus on the interrelations among different category labels, we have implemented the CFFM. This module encompasses domain knowledge embedding, the entity relationship network (ERN), and a mechanism for extracting label relevance. Detailed elaborations of these components will ensue.

3.3.1. Domain knowledge embedding

By transforming the TCM symptom ontology framework into a knowledge graph, we represent it as a graph $G = (V, E)$. The core idea of the graph convolutional network (GCN) [44] is to update the feature representation of each node by aggregating information from its neighboring nodes. For each symptom node v_i in the graph, with a feature representation h_i , we calculate the updated feature representation h'_i using the following formula:

$$h'_i = \sum_{j \in N(v_i)} \frac{1}{c_{ij}} \cdot W \cdot h_j \quad (2)$$

where $N(v_i)$ represents the set of neighboring nodes of node v_i , c_{ij} is the normalization factor, typically the number of neighboring nodes of node v_i , and W is the weight matrix used for linearly transforming the feature representations of neighboring nodes. We update the feature representation of each node through iterations until convergence, thus ultimately obtaining the embedding representation H_g .

3.3.2. Entity relationship network

The objective of this module is to generate characteristic representations of factual information and is primarily composed of the biaffine mechanism [45], attention, and MLP. First, we concatenate the entity embedding $v_i \in \mathbb{R}^{d_v}$ and position embedding $p_i \in \mathbb{R}^{d_p}$ to obtain the following comprehensive representation of the entity:

$$e_i^* = [v_i; p_i] \in \mathbb{R}^{d_v+d_p} \quad (3)$$

Next, we use the biaffine mechanism to calculate the correlation scores between entities, denoted as z_{ij} :

$$z_{ij} = \text{biaffine}(e_i^*, e_j^*) = e_i^* \cdot W \cdot e_j^* \quad (4)$$

where z_{ij} represents the correlation score between the entity e_i and the entity e_j , and W represents the learnable parameter matrix.

Next, we use the attention mechanism to compute the attention weight distribution of each entity towards the entire text sequence C , denoted as a_i , where $C = [H_e; H_g]$:

$$a_i = \text{Attention}(e_i^*, C) \quad (5)$$

where $a_i \in \mathbb{R}^T$ represents the attention weights of the entity e_i towards each position in the text sequence. The attention mechanism allows for adaptive focusing on different parts of the text sequence based on the entity's information.

Finally, we further process the obtained correlation score z_{ij} and attention weight distribution

a_i using an MLP:

$$r' = \text{MLP}(z_{ij}, a_i) \quad (6)$$

where r' represents the final correlation score between the entity e_i and the entity e_j .

3.3.3. Label correlation extraction

We calculate the attention weights between labels using the following formula:

$$e_{ij} = \text{ReLU}(W_1 \cdot \text{concat}(L_i, L_j) + b_1) \quad (7)$$

where L_i is the embedding vector of label i , L_j is the embedding vector of label j , $\text{concat}(L_i, L_j)$ represents concatenating these two embedding vectors, W_1 and b_1 are the weights and bias parameters of the MLP, respectively, and ReLU is the activation function. Note that this is only to calculate attention scores, and further normalization is needed.

To do this, we use the softmax function to normalize the attention scores and obtain the attention weights a_{ij} between label i and label j :

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^N \exp(e_{ik})} \quad (8)$$

Finally, we use the attention weights to calculate the correlation score r'' between label i and label j :

$$r'' = a_{ij} \cdot (K_i \cdot K_j^T) \quad (9)$$

where K_i and K_j are the knowledge graph embedding vectors for label i and label j , respectively, and $K_i \cdot K_j^T$ represents their dot product. After obtaining the entity correlation features and the label features, we use the biaffine mechanism to obtain the relevance scores S between the input embedding H_e and all correlation features $R = [r'; r'']$:

$$S = H_e U R + b_2 \quad (10)$$

where S is the relevance score matrix obtained from a bilinear transformation, thereby correlating the input entity's embedding H_e with its associated features in R . H_e encapsulates the entity's characteristics within the ontology, while U modulates the interaction between H_e and R , thus capturing their associative dynamics. The bias vector b_2 introduces an offset to fine-tune the model's activation levels, thereby enhancing the alignment of relevance scores with the semantic context.

3.4. Multi-label classification module

To tackle the challenge of label imbalance in multi-label classification tasks, we have adopted the

Cascade-XML model [42], which is guided by two principal considerations. First, Cascade-XML leverages a multi-tier resolution learning pipeline and a hierarchical label tree, thus effectively alleviating the label imbalance and boosting the classification accuracy for less-represented categories. Second, in the context of the complex corpora, including TCM diagnostic texts, the employment of the CFFM effectively captures the interrelationships between labels. This approach significantly diminishes the dependency on annotated datasets and concurrently augments the clustering efficacy of the HLT.

Cascade-XML is underpinned by three pivotal components: 1) the employment of the pre-trained language model BERT to extract text representations across multiple levels and its integration with the BERT + Bi-LSTM + CRF entity recognition framework. This synergy maximizes BERT's linguistic representation capabilities, thereby significantly enhancing the model's overall performance; 2) the introduction of a hierarchical label tree, which meticulously refines the clustering of label spaces, thereby establishing an increasingly detailed hierarchical structure; and 3) a set of linear classifiers $W_{(t)}$, with each meta-classifier selecting a specific level a to extract features at a particular resolution.

The final entity classification is achieved by computing scores $(w_l, \phi(x))$ for each label using label-aware weight vectors w_l .

At each level of the tree hierarchy, the objective is to precisely recognize the most likely meta-label; this can be accomplished by minimizing the one-minus-all loss:

$$L^{(t)}(x, y) = \frac{1}{|S^{(t)}|} \sum_{l \in S^{(t)}} L^{(t)}(\langle w_{kl}^{(t)}, v^{(t)}(x) \rangle, R^{(t)}(y)_l) \quad (11)$$

where $S^{(t)}$ represents the label filtering space, $v_{(x)}^{(t)}$ denotes the corresponding classification features, and $R^{(t)}(y)_l$ signifies the label set for each layer.

Therefore, the overarching objective is as follows:

$$L_{class} = \sum_{t=1}^{T+1} \alpha^{(t)} L^t(x, y) \quad (12)$$

where $\alpha^{(t)} = |S^{(t)}| / \min_{t \in [T+1]}(|S^{(t)}|)$ is used for rescaling the losses across multiple resolutions.

Lastly, we collectively optimize the objectives during the training phase as follows:

$$L_{total} = \lambda L_{entity} + (1 - \lambda) L_{class} \quad (13)$$

We introduce the task weight parameter λ to modulate the contributions of the two tasks to the loss. During the model training process, we minimize the aforementioned loss function through back-propagation and the Adam optimizer, thus iteratively adjusting the model parameters to steadily enhance its performance in the multi-label classification task.

4. Experiments

4.1. Datasets

In this study, we meticulously constructed an experimental dataset by integrating a comprehensive symptom ontology framework derived from TCM. This dataset was developed systematically,

combining rule-based methodologies, manual annotations, and subsequent corrections. Specifically, we selected anonymized data from the Qi-Huang TCM electronic medical records, thus aligning our methodology with Yang et al. [46]. The dataset encompasses a substantial collection of 35,355 electronic medical records rich in TCM diagnostic information, including symptoms, syndromes, and various physical examination details. Detailed data information is shown in Table 2.

Table 2. Statistics for the dataset used in the experiment.

Category	Average Text Length	Number of Entities	Number of Labels	Number of Texts
Western Medicine Diagnosis	9	12,994	7	6188
Present Illness History	310	236,445	46	31,526
Patient's Complaint	105	73,834	29	13,931
Chief Complaint	10	88,511	9	23,922
Inspection Diagnosis	12	29,461	11	10,522
Pulse Diagnosis	18	117,982	9	30,252
Tongue Diagnosis	15	109,162	16	30,323
Physical Examination	9	22,524	6	16,089
Total	-	690,913	133	162,753

Our dataset selectively incorporated specific textual segments from these records. These segments encompassed a range of categories such as 'Western Medicine Diagnosis,' 'Present Illness History,' 'Patient's Complaint,' 'Chief Complaint,' 'Inspection Diagnosis,' 'Pulse Diagnosis,' 'Tongue Diagnosis,' and 'Physical Examination.' Altogether, this process resulted in extracting 690,913 symptom entities for experimental purposes. In our pursuit of optimizing the model's efficacy, we undertook a rigorous preprocessing routine. This involved the removal of stopwords and special characters tailored to accommodate the variable lengths of different text types. To maintain consistency, we truncated texts exceeding 400 words, such as those in the 'Present Illness History' segment, which typically detail the patient's current symptoms, symptom progression, and treatments received. Conversely, we employed a concatenation approach for shorter texts such as 'Inspection Diagnosis' and 'Pulse Diagnosis,' thereby averaging approximately 12 and 18 words.

The dataset was strategically partitioned using a multi-label stratified sampling technique, dividing it into an 80% training set and a 20% test set. This division ensured that each label combination was proportionately represented in both sets, thus facilitating a robust entity recognition and multi-label classification tasks.

4.2. Baseline models

To comprehensively assess our method's performance and its advantages in both entity extraction and multi-label classification sub-tasks, we conducted the following set of comparative experiments:

Bi-LSTM-CRF [47]: This model integrates the bidirectional long short-term memory (Bi-LSTM) and CRF to incorporate contextual information and distributed word representations for feature extraction. This enhances the recognition performance by maximizing the correlation between words and labels.

BERT + Bi-LSTM + CRF [48]: This model incorporates the attention mechanism of BERT, which improves the entity recognition accuracy for complex features and less evident components by leveraging pre-trained word vectors.

Attention-XML [27]: An attention-based multi-label text classification model that effectively captures any crucial information in the text and semantic associations among labels, thus leading to a more accurate multi-label classification.

SGM [49]: A multi-label classification sequence generation model that treats multi-label classification as a sequence generation problem. To address this challenge, it employs a sequence generation model with an innovative decoder structure.

Cascade-XML [42]: An end-to-end multi-resolution learning pipeline that utilizes the multi-layer structure of transformer models to focus on various label resolutions, thus employing independent feature representations for an optimal label subset selection.

4.3. Experimental setup

To ensure the maximal efficiency and circumvent the risk of overtraining, we designed the training procedure to terminate if no improvement in the model was observed for more than 40 consecutive checkpoints. The model development was executed using Python 3.8 within the PyCharm environment. We utilized the Adam optimization algorithm to tackle gradient challenges inherent in NLP tasks. A batch size of 100 was employed for model training, and the process was conducted on a robust computing system equipped with a 24-core RTX 3090 GPU. The training regimen encompassed 50 iterations, thus establishing strategic checkpoints every 20 batches to secure the optimal model.

4.4. Experimental results and analysis

4.4.1. Comparison with baseline models

Our research benchmarked the proposed model against five foundational models, specifically focusing on its efficacy in classifying symptom labels within an extensive label space. As detailed in Table 3, the results reveal that the multi-label classification methodology exhibits substantial superiority over conventional entity recognition approaches, such as Bi-LSTM-CRF and BERT + Bi-LSTM + CRF. Notably, in comparison to the BERT + Bi-LSTM + CRF model, the Attention-XML and SGM, which is integral to our multi-label classification strategy, yielded an increase of 2.51 and 3.13% in Micro-F1 scores, respectively, coupled with a decrease of 0.58 and 1.77% in the Hamming Loss, respectively. This optimization can be ascribed to the traditional models' challenge in managing intricate label spaces characterized by many label categories and complex interdependencies. Through the construction of HLT, our approach effectively navigates these challenges.

Moreover, compared to the contemporary Cascade-XML model, our model registers considerable enhancements across all evaluative metrics. There was a notable reduction in the Hamming Loss by 0.39% and an increase in the Micro-F1 score by 1.6%. This marked performance improvement is attributed to the novel design of our model, which synergizes an ontological framework with the CFFM. This hybridization facilitates the extraction of pertinent entity features and adeptly captures the intricate relationships among labels. Furthermore, applying advanced multi-label classification techniques for processing entity features endows our model with a heightened efficiency and accuracy, particularly

in multi-label processing within TCM symptom texts.

Table 3. Comparison of results between baseline models and our proposed approach.

Model	Hamming Loss ($*10^{-2}$)	Accuracy	Recall	Micro-F1
Bi-LSTM-CRF	6.790	0.7712	0.7636	0.7737
BERT + Bi-LSTM + CRF	6.151	0.7954	0.7834	0.7852
Attention-XML	5.571	0.8175	0.8032	0.8103
SGM	4.381	0.8236	0.8094	0.8165
Cascade-XML	3.329	0.8301	0.8282	0.8292
Proposed	2.932	0.8388	0.8325	0.8452

4.4.2. Ablation analysis

In this section, we conducted ablation analyses to assess the contributions of various components within our model. Each experiment was configured as follows: 1) -w/o MER: excluding the entity recognition component; 2) -w/o CFFM: omitting the correlation feature fusion module; and 3) -w/o HLT: not utilizing the hierarchical label trees. Table 4 presents a summary of the performance results on the test dataset.

Table 4. Results of ablation analyses for different components.

Model	Micro-F1	Hamming Loss($*10^{-2}$)
Proposed	0.8352	3.132
-w/o MER	0.7903 (-0.0449)	6.148 (+3.016)
-w/o CFFM	0.8124 (-0.0228)	4.369 (+1.437)
-w/o HLT	0.8098 (-0.0354)	5.534 (+2.402)

A series of ablation studies were performed to rigorously assess the individual model components' contribution. Removal of the NER component led to a notable decrement of 4.49% in the Micro-F1 score and an increment of 3.016% in the Hamming Loss, thus underscoring the NER component's critical role in crucial entity extraction and its integral impact on the overall model efficacy. Further analysis indicated that excluding the CFFM precipitated a 2.28% decline in the Micro-F1 and a 1.437% rise in the Hamming Loss, thus highlighting the CFFM's vital function in the entity feature processing and in augmenting the handling of complex label interrelations. Lastly, the non-utilization of the HLT resulted in a 3.54% reduction in the Micro-F1 and a 2.402% increase in the Hamming Loss, thus emphasizing the HLT's significance in label classification management and in addressing label imbalance challenges.

4.4.3. Class imbalance

Upon analyzing the dataset, it was observed that prevalent symptom labels constituted 57% of the total samples, whereas rare labels comprised a mere 0.1%. The imbalance in label distribution is illustrated in Figure 3, thus highlighting a potential challenge for model training, especially in predicting infrequent symptoms.

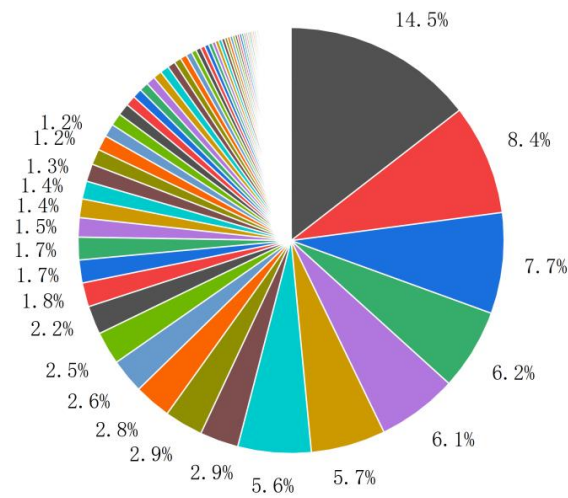


Figure 3. Visualization of label imbalance results.

A comprehensive set of experiments evaluated the model's efficacy in addressing data imbalances and forecasting low-frequency class labels (see Figure 4). The findings revealed an increase in the Hamming Loss and a decrease in the F1 score following the removal of high-frequency labels. Without a HLT, the Hamming Loss escalated from 3.534 to 7.25, while the F1 score diminished from 0.8298 to 0.6698. Conversely, implementing HLT resulted in more consistent model performance, with the Hamming Loss reaching 6.157 and the F1 score reducing to 0.6925. These results underscore the pivotal role of HLT in enhancing label hierarchy management. By establishing a structured hierarchical label framework, HLT significantly improved the model's responsiveness to minority class labels and mitigated the disproportionate impact of standard class labels on classification outcomes.

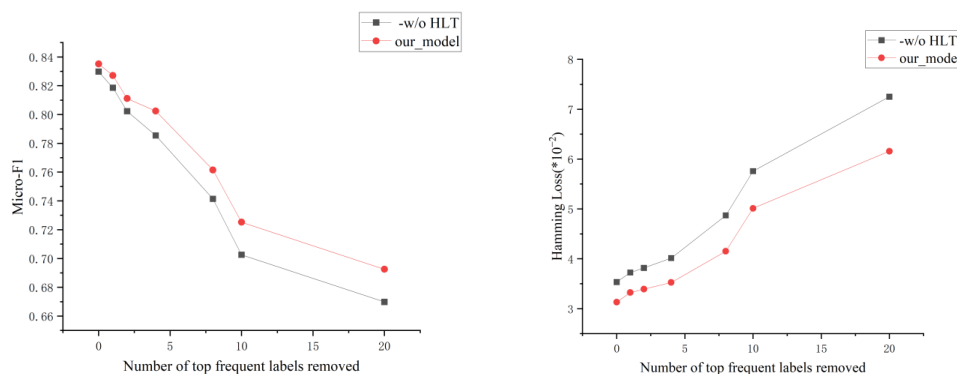


Figure 4. The Impact of High-Frequency Labels on Model Performance. This chart illustrates the change in model performance as the most frequent labels are removed incrementally. The horizontal axis represents the number of high-frequency labels removed.

In this study, we meticulously selected low-frequency labels representing varying proportions

within our data set and conducted a detailed visualization of their performance outcomes, as depicted in Figure 5. Upon a rigorous analysis, it became evident that implementing a HLT surpassed the efficacy of models that did not markedly incorporate HLT, particularly across all the minority category labels under examination. This was especially pronounced in categories with a shallow representation. For instance, in the “Symptoms of the Voice” category, which accounts for a mere 0.1% of the data, the application of HLT facilitated a substantial enhancement in the F1 score, elevating it from 0.445 to 0.615. This significant improvement underscores the HLT structure’s profound impact on augmenting the category recognition capabilities with a minimal occurrence.

Furthermore, the HLT approach demonstrated its proficiency in strengthening the model performance across other diverse categories, such as “Symptoms of the Limbs,” “Symptoms of the Diet and Taste,” and “Body Constitution.” These findings robustly endorse the HLT’s utility in effectively balancing the distribution of categories within multi-label datasets. The HLT’s particular effectiveness in elevating the classification precision of the minority categories emerges as a pivotal aspect of our study, thus underscoring its potential for broader applications in multi-label data classification tasks.

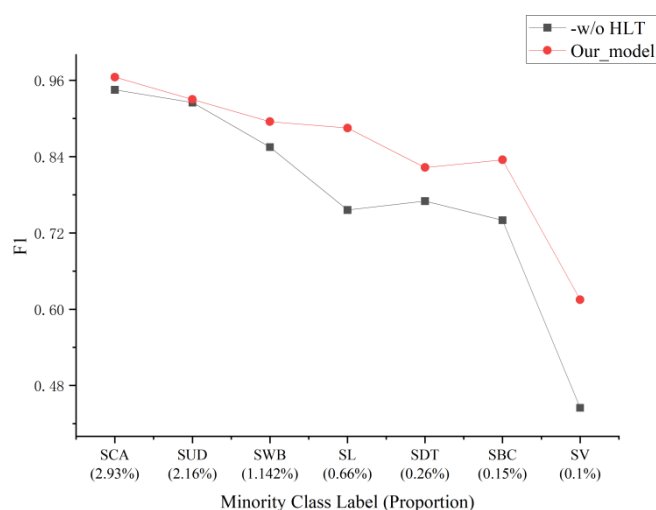


Figure 5. Performance Comparison of Low-Frequency Sample Labels. The labels include SCA (Symptoms of the Chest and Abdomen), SUD (Symptoms of Urination and Defecation), SWB (Symptoms of the Whole Body), SL (Symptoms of the Limbs), SDT (Symptoms of the Diet and Taste), SBC (Symptoms of the Body Constitution), and SV (Symptoms of the Voice).

4.4.4. The influence of label quantity on model performance

In our research, we explored the impact of label quantity on the performance of BERT + Bi-LSTM + CRF (BBC), Cascade-XML, and our proposed model, as depicted in Figure 6. The data revealed that all three models exhibited performance fluctuations with an increase in the number of labels. The BBC model demonstrated a performance decrease from 0.83 with a single label to 0.78 with the full spectrum of labels. The Cascade-XML model exhibited minimal performance variations, declining from 0.8388 to 0.8299, though essentially maintaining a level above 0.83. Notably, our model consistently surpassed the performance of these two models across various label counts. It achieved

an optimal performance of 0.8451 with a single label and sustained performance of 0.8379 with all labels, thus indicating a high degree of robustness to the escalation in label numbers.

This stability is attributed to the strategic design of our model, which includes the implementation of an HLT for the efficient label management and the integration of a CFFM. The HLT effectively mitigates the complexity of handling numerous labels through its hierarchical structure. Concurrently, by exploiting the correlations between labels, the CFFM enhances the model's accuracy in recognizing and categorizing entities related to specific labels, especially in multi-label scenarios.

It is essential to highlight that during the expansion of label sets, our model and Cascade-XML demonstrated a notable performance pattern characterized by an initial decline followed by a subsequent increase. This trend may be attributed to the model's initial phase of adjusting to the challenges posed by minority categories within an imbalanced class distribution, which initially leads to decreased performance metrics. However, as the label set expands, the model likely develops and adapts new strategies to effectively address this imbalance by incorporating a more significant number of high-frequency labels. Consequently, this adaptation results in a notable recovery and potential enhancement in the model's overall performance. This observation underscores the dynamic nature of the model's learning process in response to evolving data landscapes and the complexity inherent in managing class imbalances.

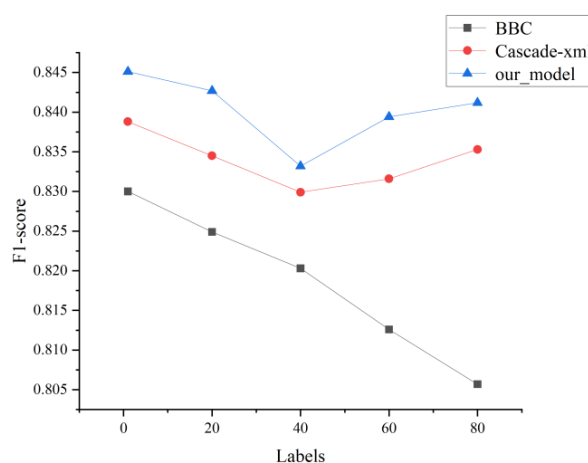


Figure 6. The impact of the number of labels on model performance.

4.4.5. Correlation

In our research, we deployed a multi-label classification algorithm that integrates an HLT based on quantitatively assessed correlations between labels. The CFFM embedded within the model is tailored to the extract label features from an ontological framework more effectively. A comparative analysis evaluated our model's performance in capturing label relationships and constructing HLT against the TF-IDF-based Cascade-XML model, with label correlations visualized through heatmap luminance levels.

As illustrated in Figure 7, results indicate that our model exhibits an enhanced sensitivity in detecting inter-label correlations compared to the Cascade-XML model. The heatmap highlights several highly correlated label pairs and groups with correlation scores around 0.5, thus demonstrating

our approach's feature extraction efficiency. In contrast, the Cascade-XML model's correlation scores were generally below 0.2 in processing complex label relationships within TCM texts. Our model significantly improved the performance in multi-label classification tasks by identifying and enhancing the training process with strongly correlated label pairs. This method increased the model's precision in recognizing symptomatic entities and their latent connections within intricate TCM texts. Moreover, utilizing this visualization technique afforded a deeper understanding of the model's operational characteristics when handling specific TCM texts, thereby enabling us to refine its architecture and training methodologies and augment its overall performance.

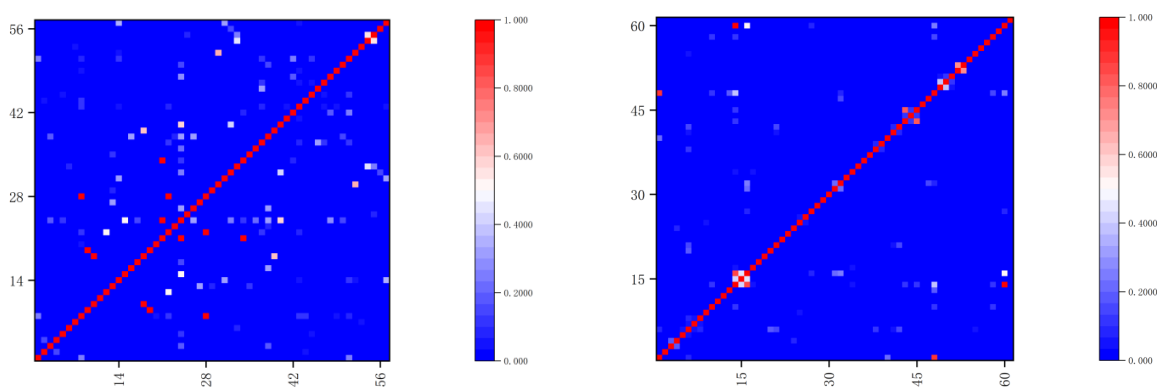


Figure 7. Label Correlation Visualization: On the left, the visualization of label correlations driven by knowledge is presented, while on the right, the data-driven visualization results are displayed.

4.4.6. Knowledge fusion analysis

A comprehensive comparative experiment was undertaken to ascertain the efficacy of the developed ontology framework. This experiment scrutinized the performance of three distinct models when applied to a uniform dataset of TCM symptoms: our model, which integrates the ontology framework; a model devoid of the ontology framework (-w/o Onto); and the conventional BBC model. The empirical analysis drew upon authentic case texts from Qihuang TCM electronic medical records, as depicted in Table 5. In the entity recognition comparative experiment conducted on the Qihuang TCM Electronic Medical Records dataset, the ontology framework integrated model (i.e., our model) demonstrated a pronounced superiority in detecting symptom entities. According to the statistical analysis, our model successfully identified all 13 listed symptom entities, whereas the model without an ontology framework integration (-w/o Onto) recognized 11, and the BBC model identified only seven. Notably, symptoms such as “Cold feet” and “Well-formed bowel movements” were exclusively detected by our model and did not appear in the recognition lists of the other two models, thus highlighting the significant role of the ontology framework in finely differentiating symptom entities. Moreover, the recognition of symptoms “Non-yellow urine” and “Difficulty in falling asleep” by our model was not observed in the BBC model, further corroborating the value of the ontology framework in capturing nuances and enhancing the comprehensiveness of the model. Overall, the experimental

outcomes conclusively affirm the significant advantages of the ontology framework in elevating the precision and completeness of the TCM symptom entity identification.

Table 5. Case study analysis and model recognition outcomes.

Case Example: Following postoperative treatment for urolithiasis, the patient reported a notable alleviation of dizziness, although slight vertigo persisted after movements. Sensitivity to cold and wind decreased slightly compared to before. However, the patient still experienced cold feet and lower limb coldness extending to the knees, occasionally necessitating the wearing of cotton trousers. The patient struggled to fall asleep, experienced frequent dreaming, and was easily awakened. Mild fatigue ensued after minimal exertion. No significant dry mouth was reported, and there was an average appetite; the patient experienced intermittent buzzing in the head. A sensation of comfort in the epigastric region was noted after consuming cold food, non-yellow urine, well-formed bowel movements, and Suboptimal mood.

Symptomatic Entity	our_model	our_model -w/o Onto	BBC
Dizziness	√	√	√
Sensitivity to cold	√	√	√
Sensitivity to wind	√		
Cold feet	√	√	
Lower limb coldness	√	√	√
Struggled to fall asleep	√	√	
Frequent dreaming	√	√	√
Easily awakened	√		
Dry mouth	√	√	√
Average appetite	√	√	
Non-yellow urine:	√	√	
Well-formed bowel movements	√	√	√
Suboptimal mood	√	√	√

Additionally, we assessed the efficacy of three models in fine-grained label recognition, as depicted in Figure 8. The traditional BBC model displayed the lowest performance in recall, precision, and F1 scores, thus indicating its limited capability in accurately identifying fine-grained labels within the TCM symptom dataset. In contrast, the model lacking the ontology framework (-w/o Onto) showed some improvement; however, the ontology-integrated model (i.e., our model) outperformed it. Our ontology-integrated model demonstrated superior performance across the recall, precision, and F1 scores. It achieved exceptionally high or exact recall rates of one for specific labels, thus underscoring its precision and comprehensiveness in identifying symptom labels in real-case texts.

Our ontology-integrated model showed a notable advantage in processing the TCM symptom dataset. By incorporating the ontology framework, with its structured symptom classification and relational network, we significantly enhanced the model's precision in identifying and classifying symptom entities, especially those with complex semantics and nuanced distinctions. Moreover, the ontology framework strengthened our model's ability to discern relationships between symptoms frequently encountered in TCM texts. This improvement was crucial for handling the combinations of symptoms typically found in TCM, thereby enhancing the model's effectiveness in multi-label classification tasks.

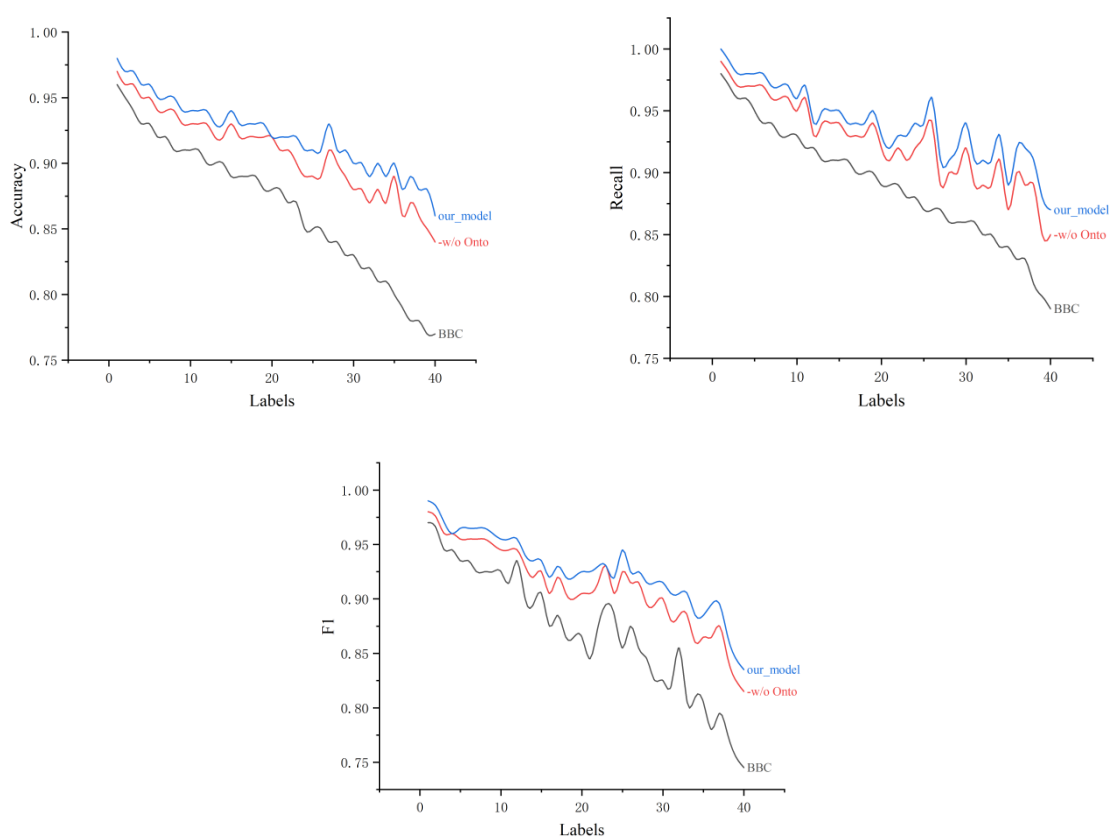


Figure 8. Model performance analysis: Fine-grained label recognition metrics including accuracy, recall, and F1 score.

4.5. Discussion

Ontology framework and standardization of TCM symptom nomenclature. Our study established a TCM symptom ontology derived from authoritative texts, thereby effectively addressing the non-standardization challenge in TCM symptom terminology. This ontology significantly streamlines the dataset construction and annotation by providing a standardized symptom classification and a description framework.

Integration of TCM knowledge with advanced modeling techniques. We bridged the gap in the existing studies regarding the integration of TCM knowledge. We significantly enhanced the model performance and annotation quality by translating our ontology into a knowledge graph and incorporating it into the model. Integrating this TCM expertise is crucial for improving symptom identification, classification accuracy, and efficacy.

Enhanced symptom entity identification with CFFM. Our approach effectively overcomes the limitations of the current sequence labeling models in complex multi-label classification tasks. The CFFM we developed combines multi-label and entity recognition processes, thus optimizing symptom entity identification. This model demonstrated a superior performance in capturing extensive label correlation features and establishing a hierarchical label system.

Comparative analysis of model performance. Regarding model performance, while the BBC model showed proficiency with datasets with fewer label categories, its effectiveness decreased with increasing labels. In contrast, our model and Cascade-XML exhibited enhanced adaptabilities and robustness with larger label counts. This is attributed to their structured outputs and the ability to capture inter-label dependencies, thus effectively managing multi-label challenges.

Limitations and future research directions. Despite these advancements, our study faces limitations. The primary constraint is the limited scope of the dataset, which was sourced from a single hospital and might affect the generalization of our model. Future research should focus on validating our model across more diverse datasets. Additionally, our study predominantly concentrated on symptom identification and classification, thus excluding treatment recommendations - a vital aspect of TCM diagnostics. This presents a promising avenue for future exploration. Moreover, while our ontology framework lays a structured foundation for symptom information, it is essential to enhance the depiction of symptom attributes for a more comprehensive understanding of TCM symptomatology. Future efforts will incorporate heterogeneous datasets and refine the ontology framework, thereby improving the accuracy of TCM clinical pattern differentiation and contributing more effectively to TCM diagnostics.

5. Conclusions

We present an innovative multi-label entity extraction model dedicated to symptom identification and classification, underpinned by a thorough TCM symptom ontology framework. This model synergizes the strengths of the TCM symptom ontology, knowledge graphs, classical entity recognition, and advanced multi-label classification techniques. A sophisticated multi-associative feature fusion module within the model significantly enhances the discernment of interconnections among symptom entities in textual data, thereby augmenting its ability to efficiently extract and interpret textual information. This integrative approach boosts the model's overall efficacy and precision and aids in the meticulous recognition, categorization, and progressive evolution of symptom entities. There are plans to merge this framework with the ICPC3 Level-3 Directory and the OHDSI conceptual structure framework. This strategic alignment aims to bridge the gap between Eastern and Western medical practices and to continuously polish the TCM symptom ontology framework for broader applicability and impact.

Use of AI tools declaration

This research was not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the Jiangxi Provincial Natural Science Foundation (20224BAB206102), National Natural Science Foundation of China (82260988), Scientific and Technological Research Project of Jiangxi Provincial Department of Education (GJJ2200923) and Science and Technology Planning Project of Jiangxi Provincial Health and Family Planning Commission (202211404).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Q. Jia, D. Zhang, S. Yang, C. Xia, Y. Shi, H. Tao, et al., Traditional Chinese medicine symptom normalization approach leveraging hierarchical semantic information and text matching with attention mechanism, *J. Biomed. Inf.*, **116** (2021), 103718. <https://doi.org/10.1016/j.jbi.2021.103718>
2. Z. Huang, J. Miao, J. Chen, Y. Zhong, S. Yang, Y. Ma, et al., A traditional Chinese medicine syndrome classification model based on cross-feature generation by convolution neural network: model development and validation, *JMIR Med. Inf.*, **10** (2022), e29290. <https://doi.org/10.2196/29290>
3. T. Bai, H. Guan, S. Wang, Y. Wang, L. Huang, Traditional Chinese medicine entity relation extraction based on CNN with segment attention, *Neural Comput. Appl.*, **34** (2022), 2739–2748. <https://doi.org/10.1007/s00521-021-05897-9>
4. A. Roy, S. Pan, Incorporating medical knowledge in BERT for clinical relation extraction, in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, (2021), 5357–5366. <https://aclanthology.org/2021.emnlp-main.435>
5. P. Chandak, K. Huang, M. Zitnik, Building a knowledge graph to enable precision medicine, *Sci. Data*, **10** (2023), 67. <https://doi.org/10.1038/s41597-023-01960-3>
6. G. Zhou, E. Haihong, Z. Kuang, L. Tan, X. Xie, J. Li, et al., Clinical decision support system for hypertension medication based on knowledge graph, *Comput. Methods Programs Biomed.*, **227** (2022). <https://doi.org/10.1016/j.cmpb.2022.107220>
7. D. Zhang, Q. Jia, S. Yang, X. Han, C. Xu, X. Liu, et al., Traditional Chinese medicine automated diagnosis based on knowledge graph reasoning, *Comput. Mater. Contin.*, **71** (2022). <https://doi.org/10.32604/cmc.2022.017295>
8. Y. An, X. Xia, X. Chen, F. X. Wu, J. Wang, Chinese clinical named entity recognition via multi-head self-attention based Bi-LSTM-CRF, *Artif. Intell. Med.*, **127** (2022), 102282. <https://doi.org/10.1016/j.artmed.2022.102282>
9. Y. Ma, Y. Liu, D. Zhang, J. Zhang, H. Liu, Y. Xie, A multigranularity text driven named entity recognition CGAN model for traditional Chinese medicine literatures, *Comput. Intell. Neurosci.*, **2022** (2022), 1495841.
10. R. Qi, P. Lv, Q. Zhang, M. Wu, Research on Chinese medical entity recognition based on multi-neural network fusion and improved tri-training algorithm, *Appl. Sci.*, **12** (2022), 8539. <https://doi.org/10.3390/app12178539>
11. Y. Li, X. Wang, L. Hui, L. Zou, H. Li, L. Xu, et al., Chinese clinical named entity recognition in electronic medical records: Development of a lattice long short-term memory model with contextualized character representations, *JMIR Med. Inf.*, **8** (2020), e19848.
12. M. Zhang, J. Wang, X. Zhang, Using a pre-trained language model for medical named entity extraction in Chinese clinic text, in *2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, *IEEE*, (2020), 312–317. <https://doi.org/10.1109/ICEIEC49280.2020.9152257>

13. R. Xie, Y. Wang, D. Peng, X. Liu, X. Su, X. Li, Research on standardization of traditional Chinese medicine symptoms, *Henan Tradit. Chin. Med.*, **7** (2017), 11441146. <https://doi.org/10.16367/j.issn.1003-5028.2017.07.0403>
14. K. Zhou, J. Dong, S. Wang, G. Li, Y. Zheng, T. Wang, A review of research ideas and methods for standardization of traditional Chinese medicine symptoms in the past 20 years, *Glob. Tradit. Chin. Med.*, **4** (2022), 708–712.
15. Y. J. Hui, Q. L. Zha, A review of traditional Chinese medicine symptom information extraction, *Comput. Eng. Appl.*, **59** (2023), 35–47.
16. L. Ma, J. Li, The significance and methodology of standardization of symptom nomenclature, *Liaoning J. Tradit. Chin. Med.*, **37** (2010), 1264–1265.
17. W. Liu, F. Zhu, Reflections on several issues in the standardization of traditional Chinese medicine symptoms, *J. Tradit. Chin. Med.*, **48** (2007), 555–556.
18. D. Yan, M. Cui, Exploration of ‘symptoms and signs’ classification in the traditional Chinese medicine clinical terminology system, *Chin. J. Med. Libr. Inf. Sci.*, **10** (2015), 77–80.
19. Y. Dong, M. Cui, Discussion on the classification of ‘symptoms and signs’ in the clinical terminology system of traditional Chinese medicine, *Chin. J. Med. Libr. Inf. Sci.*, **24** (2015), 77–80.
20. X. Ling, D. Weld, Fine-grained entity recognition, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **26** (2012), 94–100. <https://doi.org/10.1609/aaai.v26i1.8122>
21. K. Pu, H. Liu, Y. Yang, W. Lv, J. Li, Multi-label fine-grained entity typing for baidu Wikipedia based on pre-trained model, in *China Conference on Knowledge Graph and Semantic Computing*, **1553** (2021), 114–123. https://doi.org/10.1007/978-981-19-0713-5_13
22. X. Ren, W. He, M. Qu, Label noise reduction in entity typing by heterogeneous partial-label embedding, in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, (2016), 1825–1834. <https://doi.org/10.1177/0165551521998048>
23. Y. Onoe, G. Durrett, Learning to denoise distantly-labeled data for entity typing, preprint, arXiv:1905.01566. <https://doi.org/10.48550/arXiv.1905.01566>
24. H. Zhang, D. Long, G. Xu, M. Zhu, P. Xie, F. Huang, et al., Learning with noise: improving distantly-supervised fine-grained entity typing via automatic relabeling, in *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, (2021), 3808–3815. <https://doi.org/10.24963/ijcai.2020/527>
25. M. A. Ali, Y. Sun, B. Li, W. Wang, Fine-grained named entity typing over distantly supervised data based on refined representations, in *Proceedings of the AAAI Conference on Artificial Intelligence*, **34** (2020), 7391–7398. <https://doi.org/10.1609/aaai.v34i05.6234>
26. Y. Prabhu, A. Kag, S. Harsola, R. Agrawal, M. Varma, Parabel: Partitioned label trees for extreme classification with application to dynamic search advertising, in *Proceedings of the 2018 World Wide Web Conference*, (2018), 993–1002. <https://doi.org/10.1145/3178876.3185998>
27. R. You, Z. Zhang, Z. Wang, S. Dai, H. Mamitsuka, S. Zhu, Attentionxml: Label tree-based attention-aware deep model for high-performance extreme multi-label text classification, *Adv. Neural Inf. Process. Syst.*, **32** (2019). <https://arxiv.org/abs/1811.01727>
28. J. Zhang, W. C. Chang, H. F. Yu, I. Dhillon, Fast multi-resolution transformer fine-tuning for extreme multi-label text classification, *Adv. Neural Inf. Process. Syst.*, **34** (2021), 7267–7280. <https://doi.org/10.48550/arXiv.2110.00685>
29. X. Xiao, Research on Data Elements of Traditional Chinese Medicine Clinical Symptoms Based on Machine Learning, Ph.D. thesis, Hunan University of Traditional Chinese Medicine, 2018.

30. N. Zhang, X. Cao, R. Lin, B. Wang, H. Shi, H. Zhou, et al., Research on the normalization of traditional Chinese medicine symptom terms in epilepsy, *World Sci. Technol.-Modernization Tradit. Chin. Med.*, **22** (2020), <https://doi.org/10.11842/wst.20190415001>
31. M. Li, Q. Zhou, X. Luo, B. Zhu, Research on the standard and classification system of traditional Chinese medicine symptom terminology, *Chin. J. Tradit. Chin. Med. Pharm.*, **36** (2021), 4838–4842.
32. E. F. Sang, F. D. Meulder, Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition, preprint, arXiv:cs/0306050. <https://doi.org/10.48550/arXiv.cs/0306050>
33. D. Gillick, N. Lazic, K. Ganchev, J. Kirchner, D. Huynh, Context-dependent fine-grained entity type tagging, preprint, arXiv:1412.1820. <https://doi.org/10.48550/arXiv.1412.1820>
34. E. Choi, O. Levy, Y. Choi, L. Zettlemoyer, Ultra-fine entity typing, preprint, arXiv:1807.04905. <https://doi.org/10.48550/arXiv.1807.04905>
35. A. Abhishek, A. Anand, A. Awekar, Fine-grained entity type classification by jointly learning representations and label embeddings, in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, (2017), 797–807. <https://aclanthology.org/E17-1075>
36. F. López, M. Strube, A fully hyperbolic neural model for hierarchical multi-class classification, preprint, arXiv:2010.02053. <https://doi.org/10.48550/arXiv.2010.02053>
37. W. Xiong, J. Wu, D. Lei, M. Yu, S. Chang, X. Guo, et al., Imposing label-relational inductive bias for extremely fine-grained entity typing, preprint, arXiv:1903.02591. <https://doi.org/10.48550/arXiv.1903.02591>
38. Y. Fan, J. Liu, J. Tang, P. Liu, Y. Du, Learning correlation information for multi-label feature selection, *Pattern Recognit.*, 2023. <https://doi.org/10.1016/j.patcog.2023.109899>
39. Y. Fan, J. Liu, P. Liu, Y. Du, W. Lan, S. Wu, Manifold learning with structured subspace for multi-label feature selection, *Pattern Recognit.*, **120** (2021). <https://doi.org/10.1016/j.patcog.2021.108169>
40. Y. Fan, B. Chen, W. Huang, J. Liu, W. Weng, W. Lan, Multi-label feature selection based on label correlations and feature redundancy, *Knowl.-Based Syst.*, **241** (2022). <https://doi.org/10.1016/j.knosys.2022.108256>
41. Y. Fan, J. Liu, S. Wu, Exploring instance correlations with local discriminant model for multi-label feature selection, *Appl. Intell.*, **52** (2022), 1–19. <https://doi.org/10.1007/s10489-021-02799-0>
42. S. Kharbanda, A. Banerjee, E. Schultheis, R. Babbar, CascadeXML: Rethinking transformers for end-to-end multi-resolution training in extreme multi-label classification, in *Part of Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*, 2022.
43. J. Sun, F. Yang, W. Deng, Construction of a knowledge representation model for traditional Chinese medicine symptoms based on ontology, *J. Med. Inf.*, **38** (2017).
44. T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, preprint, arXiv:1609.02907. <https://doi.org/10.48550/arXiv.1609.02907>
45. J. Yu, B. Bohnet, M. Poesio, Named entity recognition as dependency parsing, preprint, arXiv:2005.07150. <https://doi.org/10.48550/arXiv.2005.07150>
46. R. Yang, Q. Ye, C. Cheng, S. Zhang, Y. Lan, J. Zou, Decision-making system for the diagnosis of syndrome based on traditional Chinese medicine knowledge graph, *Evid. Based Complementary Altern. Med.*, 2022. <https://doi.org/10.1155/2022/8693937>
47. N. Deng, H. Fu, X. Chen, Named entity recognition of traditional Chinese medicine patents based on BiLSTM-CRF, *Wireless Commun. Mob. Comput*, 2021. <https://doi.org/10.1155/2021/6696205>

48. Q. Qu, H. Kan, Y. Wu, Y. Gao, Named entity recognition of TCM text based on bert model, in *2020 7th International Forum on Electrical Engineering and Automation (IFEEA)*, 2020. <https://doi.org/10.1109/IFEEA51475.2020.00139>
49. P. Yang, X. Sun, W. Li, S. Ma, W. Wu, H. Wang, SGM: Sequence generation model for multi-label classification, preprint, arXiv:1806.04822. <https://doi.org/10.48550/arXiv.1806.04822>



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)